HFM-128/143: Moderators as a Method to Link HF to Operational Behaviour

Dr. Wouter Lotens, Ph.D.
TNO Human Factors
NLD

wouter.lotens@tno.nl

Dr. Laurel Allender, Ph.D.
U.S. Army Research Laboratory
USA

Dr. Joe Armstrong, Ph.D.
CAE Professional Services
CAN

Dr. Andrew Belyavin, PhD.
Centre for Human Sciences, QinetiQ
GBR

Mr. Brad Cain, M.A.Sc.
DRDC Toronto, Human System Integration Section
CAN

Dr. Martin Castor, Ph.D.
Swedish Defence Research Agency FOI
SWE

Dr. Kevin Gluck, PhD
AFRL
USA

Mr. Mikael Lundin, M.Sc.
Swedish Defence Research Agency FOI
SWE

ABSTRACT

In an attempt to define a generic way to include human factors in a simulation of military operations HFM 128 suggested to break the causal chain between human factors as input and operational outcome into three components: a) human factors science based models to calculate state variables of the human, b) Performance Shaping Functions (PSF) to quantify the effect of states on individual performance and c) operational models that integrate individual performance into operational performance. For a significant number of human factors state variables have been suggested that serve the intended purpose: fairly unique variables that represent the human factors states they stand for and which are in a comprehensive way related to performance. Cognitive processes are more complex in this respect, although also for this class of human factors state variables may be found. The concept of PSF has been exploited already in published performance models and gains in relevancy by using the state variables. This approach weeds out much of the multiplicity that emerges from the numerous tasks and conditions that may be considered. A compact approach to operational performance involves both (sub) task performance and the quality of the operational plan. Operational performance is then measured in success rate and military cost factors involved. None of the modelling architectures on the market exploits the full range depicted here. They seem to focus on aspects such as cognition, operational embedding, task networking or detailed behaviours. Recommendations to NATO are given to develop this field in order to cover the full needs of operational modelling, to reuse knowledge and to improve on validation of simulations.

1.0 INTRODUCTION

HFM128 reported on the feasibility of an approach to the inclusion of Human Behaviour Representation in constructive simulations (HFM 128, 2007). The key issue is that the worlds of operational analysis, human behaviour modelling and human factors sciences need to be connected to use constructive modelling as a tool for training or analysis in an operational setting. This would allow assessing the effect of a specific

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **OCT 2010**	2. REPORT TYPE **N/A**	3. DATES COVERED -

4. TITLE AND SUBTITLE **HFM-128/143: Moderators as a Method to Link HF to Operational Behaviour**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Research Laboratory USA**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

See also ADA564696. Human Modelling for Military Application (Applications militaires de la modelisation humaine). RTO-MP-HFM-202

14. ABSTRACT

In an attempt to define a generic way to include human factors in a simulation of military operations HFM 128 suggested to break the causal chain between human factors as input and operational outcome into three components: a) human factors science based models to calculate state variables of the human, b) Performance Shaping Functions (PSF) to quantify the effect of states on individual performance and c) operational models that integrate individual performance into operational performance. For a significant number of human factors state variables have been suggested that serve the intended purpose: fairly unique variables that represent the human factors states they stand for and which are in a comprehensive way related to performance. Cognitive processes are more complex in this respect, although also for this class of human factors state variables may be found. The concept of PSF has been exploited already in published performance models and gains in relevancy by using the state variables. This approach weeds out much of the multiplicity that emerges from the numerous tasks and conditions that may be considered. A compact approach to operational performance involves both (sub) task performance and the quality of the operational plan. Operational performance is then measured in success rate and military cost factors involved. None of the modelling architectures on the market exploits the full range depicted here. They seem to focus on aspects such as cognition, operational embedding, task networking or detailed behaviours. Recommendations to NATO are given to develop this field in order to cover the full needs of operational modelling, to reuse knowledge and to improve on validation of simulations.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**18**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

human factor on the individual performance, as well as on operational performance when teams work together towards a military aim. For definitions please consult Appendix 1.

Modelling is used to understand and quantify the impact of a large number of factors on the operational outcome. These factors may relate to the task, the environment, capability limitations, opponent, human factors and operational factors involved. By the large variation in conditions, empirical methods become infeasible. No one has the resources to systematically investigate the effects of these factors, in the grades in which they may apply and in the numerous interactions they may show. Models are in principle equally complex as the reality, but the understanding of the processes involved in the chain between actions of individuals and the eventual operational performance may help to comprehend the variability. It may also give a handle on tools to manage part of the chain. Figure 1 shows the breakdown of the chain as proposed by the HFM 128 Task Group. Their proposal is to define states that can be calculated or monitored. Each state is input to a tool that calculates a higher level state, until the military goal is arrived at. The performance model then consists of a number of tools (or constructions as meant in constructive modelling) that is assembled to fit the purpose.

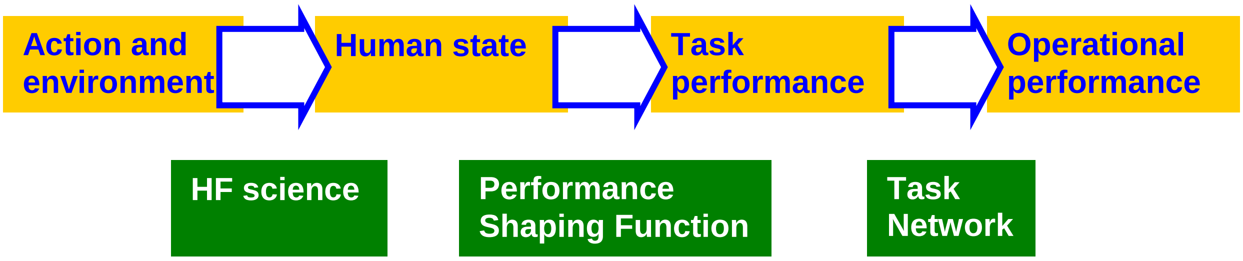

Figure 1: Breaking down the performance chain by means of three model constructional elements. Dynamic states are in yellow, constructional elements in green.

At least three constructions must be discerned in a generic performance model:

- A science based construction that is producing floating stress levels, which are regarded as critical for human performance, taking traits, treatments and interactions into account
- A performance shaping function that relates critical stress levels to task performance
- A team performance construction that takes task performance, operational activities and the tactical planning into account to arrive at operational performance and the operational costs involved

The viability of the idea must be tested. Are there scientifically sound states associated with all human factors? Is there a way to estimate the impact of states on task performance? Are Performance Shaping Functions (PSF) generic enough to be reusable for other tasks? Can the operational performance be quantified in a complex environment with many entities and an opponent with a will of its own? We will show how we envision the application of the chain in Figure 1 to the core problem of the HFM202 Symposium: to find a generic way to combine human factors and understands their impact on military task performance.

2.0 HUMAN STATES

2.1 State Variables

A human state is an in time varying parameter that is a representation of the cumulative effects of the environmental impact and task execution impacts on the human. A powerful example is body core temperature, which is the cumulative effect of weather, clothing and activity. In the ideal case a single

parameter would express all the intended impacts and nothing else. In the ideal case the parameter would also be uniquely responsible for loss of performance. That is usually not so. An elevated body temperature sets sweating in action and after some time dehydration occurs, which may become a limiting parameter, independent of body temperature. Dehydration and body temperature are not independent. However, together they cover a wide range of conditions in terms of performance. An important aspect of a state is that it allows monitoring the load on a homeostatic system, which returns to the homeostasis if the load is relieved or by behaviour (drinking). This is the way many human processes are organised and which show the adaptation (set point corrections) that makes humans so adaptive to their environments.

Figure 2 shows the generic model of a state. The mission, involving activities, and the environment cause a certain stress level, which is input to the state. Stressors (*properties of the operational environment or tasking that offset state variables from their neutral value*) change the state in a dynamical way, and traits (*characteristics of the entity that remain constant over the operation*) affect the state in a static way. The activities are performed better or worse, depending on the state of the entity. The state itself has an impact on the way activities are performed, delivering the interaction with the environment. It is important to notice that stress is an input in this context and the state is the resulting strain. Stress is the instantaneous impact on the state; the state is the accumulated effect. A useful state variable is measurable, so that it can experimentally be verified.

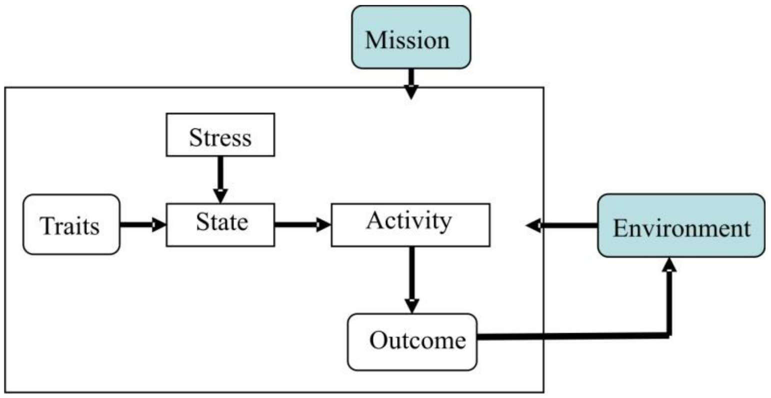

Figure 2: The HFM 128 philosophy approach to modelling operator behaviour. This approach assumes that the way activities are carried out depends on the state of the actor, resulting from his personal traits and stressors emerging from the activity itself. The state varies over time and the state variables are keys to the HF effect on performance.

State variables were found for a number of militarily relevant stressful processes. They are listed in Table 1. This table includes perceptual and physiological processes, but for task demand and threat. The concept of states seems to hold fairly well for this class a processes, but that may be different for cognitive processes, as explained in the next section.

Traits are characteristics that change only slowly or not at all, relative to the time scale of the simulation. Some traits are listed in Table 2, for individuals and for teams.

Table 1: Environmental stressors important for military simulations.

Key environmental stressors:	**State variables**
Sleep deprivation	Alertness
Rapid time zone shift – circadian effects	Alertness
Sustained physical demand	Muscle glycogen, blood glucose
Thermal effects (causing thermal strain, dehydration, discomfort)	Body temperature, dehydration
Visual environment	Detection time, identification time
Task demand - task load	Relative mental workload, error rate
Occasional environmental stressors:	
Noise (continuous and impulse)	Threshold elevation, intelligibility
Vibration	Numbness, local blood flow
Hypoxia (Loss of oxygen in high flying fast jets & work at altitude)	Blood oxygenation
Acceleration: High G for fast jets; alternating G (Push-Pull) for all aircraft	Brain blood pressure
Vestibular effects	Sickness
Threat	Anxiety

Table 2: Characteristics of individuals and groups important for military simulations.

Personal characteristics:	**Collective characteristics:**
Training (both task and physical)	Team training
Experience (on task and with equipment)	Experience (with teamwork)
Age	Team composition (Ad-hoc vs. established)
Personality, coping style and culture	Cohesion
General intelligence	Leadership
Anthropometry	Culture and Organization
Fear, Anxiety, Morale	Language

2.2 States in Cognition

Cognition, emotion and social-cultural processes are different from physiological and perceptual processes in three ways: 1) it is far more difficult to define measurable parameters, 2) the outcomes are sometimes on a nominal scale, rather than ordinal or rational and 3) cognition involves creative steps. This is most visible in decision making, where there is a discrete choice between response options that must be generated. The consequence is that architectures for cognitive modelling are completely different from physiology based models. It is thus also hard to propose a generic approach to performance modelling, facilitating the inclusion of various human factors in an integrated way.

Rather than relying on the confusing literature on the concept of Situation Awareness, HFM 143 discussed the scientific based cognitive processes, of which a summary is given in Figure 3. Although the processes are represented in an orderly way, the sequence of thoughts is usually less straightforward than depicted. Frequently jumps are made backwards when something must be verified or tried again. Also, an implicit, non-conscious evaluation is running parallel to this scheme. If the outcomes of the explicit and implicit processes do not match it is difficult to make a decision and likely the explicit process is reviewed to see how it can be bended to make the match. Bending the implicit process is not really possible and some researchers believe that decisions are always made implicitly (Wegner and Wheatley, 1999). Recently, Baumeister and Masicampo (2010) fired up this discussion in a well received paper, stating that behaviour

is controlled by unconscious processes and that conscious cognition is at best an adviser that can be put aside. The debate also received new interest from NATO with respect to the role of emotions in decision making (HFM 209). The scientific literature on this topic is developing (see for instance Leone et al, 2005).

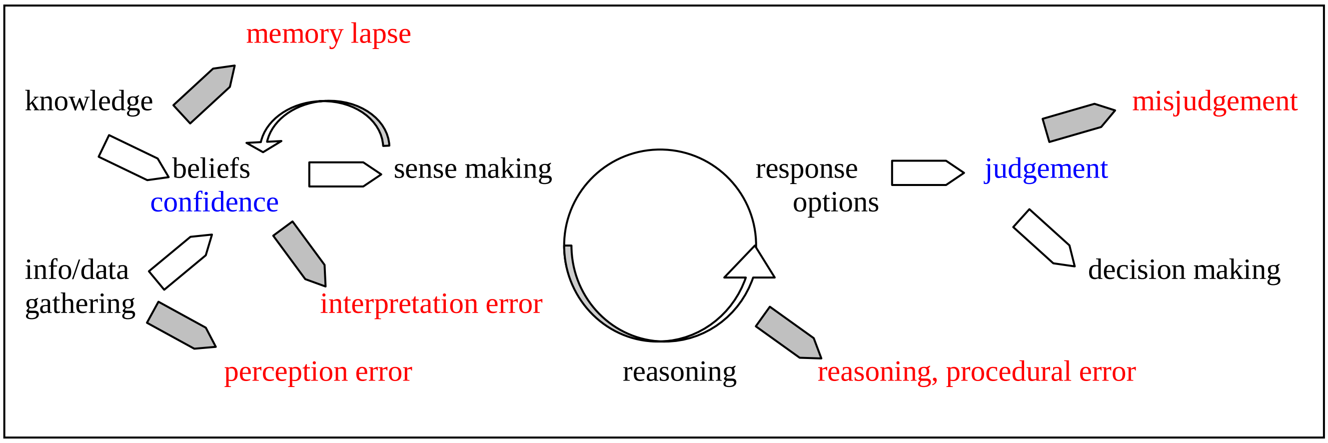

Figure 3: Mapping cognitive concepts and processes on a decision making process. The stages of Situation Awareness: information state (beliefs and confidence), understanding (sense making) and projection of future events (reasoning on response options) may be recognised and are the consequences of lower level cognitive processes. Typical error types are involved (marked in red). Concepts in blue qualify for state variables.

Decision strategies determine how the scheme is operated. For fast or effortless decisions reasoning is too cumbersome and is skipped. Instead more simple strategies are used, like relying on past experience RPDM, Klein, 1998), satisfying minimal criteria (good enough), or taking short cuts (programmed behaviour). Bratman (1999) argued in his Belief, Desire, and Intention model that efforts for decision making and execution of the decision are balanced. Effortful decision making may involve the recognition and subsequent acquisition of missing information, which indeed may take long. The evaluation is an expectation value based on all factors involved (Multi-attribute models, Sheppard et al, 1988).

Can cognition be handled as a state, similar to the stressors in Table 1? The three characteristic differences mentioned at the beginning of this section tell us that the analogy is not perfect. This is not a typical feedback control system that returns to the homeostasis. Most variables in Figure 3 tend to be stochastic and to involve discrete states, which have no order. There can not be a stress level associated with the nature of the state. However, confidence and judgment seem to be different as they evaluate the other states and rank them in order of likelihood and desirability. Confidence is interpreted here as the trust in the correctness of the hypothesis that is developed on the situation and judgment is the evaluation of a response option in terms of potential success rate to achieve the task or goal. Confidence and judgment thus would be the drivers of the cognitive process, while other mechanisms deliver the content (knowledge, beliefs, perception, sense making, reasoning and creation of response options). Confidence and judgment may qualify as state variables of the cognitive process, as satisfactory values indicate a near decision. The other variables are processes, processing content, and are unrelated to the progress of the decision making.

Frequently, researchers resort to higher order concepts to qualify cognitive states, in particular the concept of Situation Awareness. This is necessarily vaguer and less quantitative than the scientific concepts in Figure 3 and refers more to the active processes than to a state. In the discussions of HFM 128 an even broader concept was handled: the perceived world as compared to the real world. This concept was feasible to understand the impact of cognitive processes, but only in a conceptual way. It shows the desired properties of a state, but is in a concrete case too global to be useful.

Confidence and judgment may be the states monitoring the success of the cognitive process. The capacity of the brain to carry out all steps has several bottlenecks that may cause overload, error and stagnation if the brain activity increases. This is generally referred to as workload and was included in Table 1.

3.0 TASK PERFORMANCE

3.1 Modifiers

The generic concept of a modifier in model context is a factor that affects performance. If we take as an example a soldier carrying out a patrol task, he will need a certain time to complete a round, depending on numerous environmental parameters. Now, the environment changes and CBRN threat urges him to wear protective clothing. The CBRN suit will moderate his performance and to learn how, we could do a comparative experiment. Probably the experiment would include the typical ensembles worn at different alert states. Next, the temperature rises and would make another performance modifier, that also needs evaluation. Moreover, CBRN ensemble and temperature are not independent. The temperature limit comes down when a high level of protection is applied. The temperature experiment thus needs to be repeated with and without the various CBRN ensembles. This approach is depicted in Figure 4, left frame.

If the soldier is tasked with another task, the whole experimental series must be repeated. As there are also many other performance modifiers, the experimental program will soon become infeasible. Apparently, we need means to use the knowledge of one task for another, and to use the knowledge on climate and clothing to deduce the modifying effect of other climates and ensembles. This is done exploiting the states we discussed. If we would have a model that predicts the state, depending on task, clothing and climate, we could use the state as a criterion to predict task performance. The performance prediction is done by the Performance Shaping Function (PSF), depicted in Figure 4, right frame. The appropriate states in this example would be body core temperature and dehydration. By using the intermediate step of states, the complexity of performance prediction is reduced to two types of modules, a human factors science based model predicting the state and a PSF, depending on the state. The character of the PSF will be discussed below.

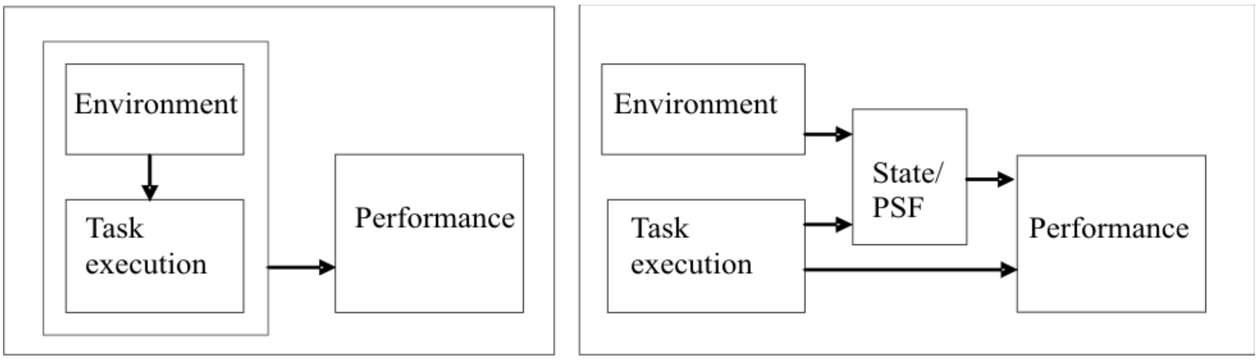

**Figure 4: Task execution leads to performance.
In the left frame the environment has a modifying effect on performance, which is heuristically established. In the right frame, the modifying effect is made explicit through states and Performance Shaping Functions. The latter functions are based on relationships between states and performance.**

3.2 The Use of States and Resources in PSF

States are a deviation of a state variable from the neutral value and usually at some level of deviation the tolerance limit is reached. The end of voluntary tolerance indicates that the underlying function is

maximally loaded or can sustain no longer. Consequently, states are indicative for continuation of a task. The simplest PSF thus looks at the states that are involved in the task and as soon as one state is at the limit, the task will no longer be performed at maximal performance. This type of PSF is already very useful, in all its simplicity. But the mechanism does not answer questions like at which lower performance level the task could be continued and how several tasks could be performed simultaneously. More detailed considerations are needed and can be found in the concept of resources. Every human function has a certain capacity, which can be used for several tasks at a time. As long as the capacity is not exceeded that particular function is not restrictive for the tasks.

Figure 5 shows how this could take effect in a model. As explained previously, the actual resources are assumed to depend on the traits (controlling the basic resources) and the states (adding dynamics to the resource). This needs to be compared to the task demands as explained above (required resource). The simplest expression for resource use would be to calculate the percentage of a particular resource that is required for a task, based on a linear, weighted sum of the resources used, and if the total demand exceeds availability, there will be some decrement. Alternatively, complex matrices have been developed to guide a modeller through the assessment of 'within task demand', 'between task demands', all depending on the nature of the resource. It is not unusual to discern perceptual, attentional, cognitive, gross motor, fine motor functions, etc. (e.g., see Wickens' Multiple Resource Theory implemented in IMPRINT (ARL, 2005)). It is not difficult to see that the linear model is deficient. When a man walks but cannot see, his performance will not drop a little, but will drop considerably, despite the fact that vision may be estimated as a minor required resource. A multiplicative model might be more realistic: performance is the multiplication of relative availabilities. Even better would be a model that knows to handle minimum and maximum requirements. The "amount" of cognition required to walk is limited but still essential since cognition is required to know where to go. Thus, a more advanced model would consider that the performance is multiplicative for all resources involved but with respective thresholds above which no further improvement may be expected.

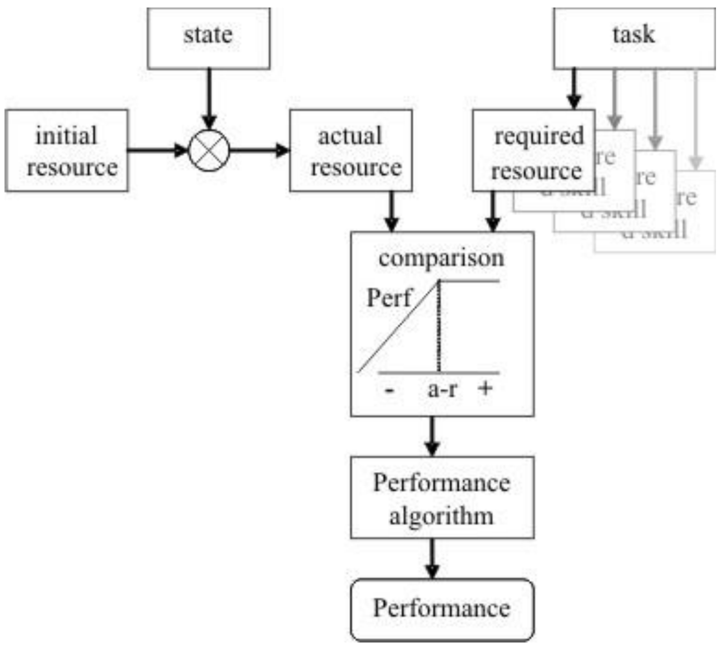

Figure 5: Task performance is moderated by resource demand and availability. A task requires resources, which may be met by the actually available resources. If so, the performance is not affected, despite the reduction of the initial resources by the current state. However, if the actual resource is less than required, performance will likely deteriorate. The Performance algorithm integrates the performance effects of the various resources involved.

3.3 Lack of Resources

One approach to resolving the resource limitation question is to make reasoned judgments about the kinds of resources available, the "amount" required by a task (i.e., task demand), and the consequence when insufficient resources are available. This is demonstrated in Figure . For example, if insufficient resources are available when a task needs to be performed, the task may be (a) performed less accurately due to perception or decision making errors; (b) performed more slowly, which adds time to the overall performance; (c) omitted, which, depending on its criticality, could be an error; or (d) delayed and performed at a later time. Or the task could be given to another human or to some automated system function.

If the capacity is exceeded by a combination of tasks, the state variable is at its maximum and the capacity will be shared over tasks is a way which may depend on priority or another strategy. Some tasks may be carried out simultaneously (e.g., parallel branches in a task network representation). However, for other types of tasks, the "calculation" of whether sufficient resources remain is more difficult. Various workload approaches exist to determine how task execution unfolds if resources are "short."

Cognitive capacity is a commonly shared resource. The question of parallel versus serial information processing in humans is an on-going debate in the literature. This debate is embodied in different cognitive architectures, for example, with EPIC arguing for no limit on central (cognitive) processing resources and ACT-R arguing for completely serial processes. Thus, a full implementation of performance moderation that links traits, states, and the environment at a fine-grained physiological and cognitive level to ultimate system performance is the goal; however, the full set of research findings to model this definitively does not exist.

Although Figure 6 does not highlight the fact that resources are not independent, of course these interdependencies should be considered. For some motor tasks, conflicts are straightforward. Tasks such as typing and playing the piano are incompatible. A task may require several resources and the actual use of resources is often higher because of interactions between resources. For example, when gross motor action is required for a task, the resource for fine motor action is reduced in a sort of parasitic way such that fine motor action is effectively disabled. An analysis of resource interactions shows that many of these interactions are asymmetrical: one hampers the other but not the reverse. Gross motor actions affect fine motor actions, perception, and cognition (for instance, particularly for high work rates and imminent exhaustion). The use of cognitive or perceptual capabilities may seem to have little effect on motor resources such as moving the whole body while walking; however, if the person is looking down at a hand-held device while walking, even the minimal perception required for walking is fully used by the other task of looking at the hand-held device. Likewise, using perceptual and cognitive resources to talk on a cell phone while driving appears to have a significant impact on attention but it may also have a significant effect on the fine motor resources required to control an automobile.

3.4 Behaviour

Although we discussed task performance, this does not really tell us what happens in reality. Tasks are performed by selecting behaviour, not by combining incidental activities. The behaviour is critical, as it is recognised as purposeful (reveals intent). Behaviour also controls the stress. Usually a certain behaviour is associated with the task and alternative behaviours could stem from attempts to counteract the effects of stress (e.g., working less due to exhaustion, moving swiftly or ducking under enemy threat, falling asleep when sleep deprived, etc.) or from reactions to events in the environment (e.g., focusing attention to the location of an adversary, fascination for a partial problem), from the consequences of carrying out ordered tasks that contain incompatible subtasks (e.g., stopping moving to aim accurately, interrupting reasoning to talk to someone), or from individual preferences. Control over behaviour consists in the first place of selection of a particular behaviour from a limited set, defined by the tasks and the alternative behaviours.

One particular challenge in task execution is the implementation of behaviour strategy. Even within tactically correct task execution, there is variability in the precise strategy employed. How would the behaviour be chosen from this possible set?

A first cut at determining which task to select is often to ask military subject matter experts (SMEs). This is often a good starting point, but it must be acknowledged that even SME judgment is limited when it comes to making estimates that can adequately populate a distribution, when new technologies or equipment are being modelled, or when millisecond-level cognitive tasks are involved. Thus, various scientific theories can be brought to bear. For example, in reinforcement learning theory, behaviour is optimized by finding strategies that balance drivers with goals where rewards for desired behaviour are balanced with the costs. The cost will be minimized to just match the rewards and since the costs include various factors (discomfort, pain, exhaustion, fear, etc.) there is also room for avoiding certain cost factors. This view on behaviour implements in a relatively straightforward way factors such as training, hardening and experience (reduction of perceived costs) as well as positive and negative motivation (manipulation of rewards by praise and punishment). Another theory proposed by Selfridge (1959) and generalized by Jackson (1987) is the concept of a Pandemonium in which "demons" are fighting for attention, each trying to "solve" the problem at hand. Dynamic behaviour is then the subsequent handling of priorities, fed by stress and task demons. A simple and task-independent algorithm would satisfy the demon that shrieks loudest, requesting a certain behaviour. Note, the COGNET architecture (Zachary, Ryder & Hicinbothom, 1998) grew from this Pandemonium notion.

At some level, the "screaming demons" could be conceptualized as the sort of level of activation notion at the heart of ACT-R (Anderson & Lebiere, 1998) such that the most active item in memory is selected for execution. This implementation is based on decades of psychological literature on learning, memory, and attention and happens within the framework of goal orientation (i.e., the idea that human behaviour is predominantly goal-directed and that a task is selected that helps to meet the current goal).

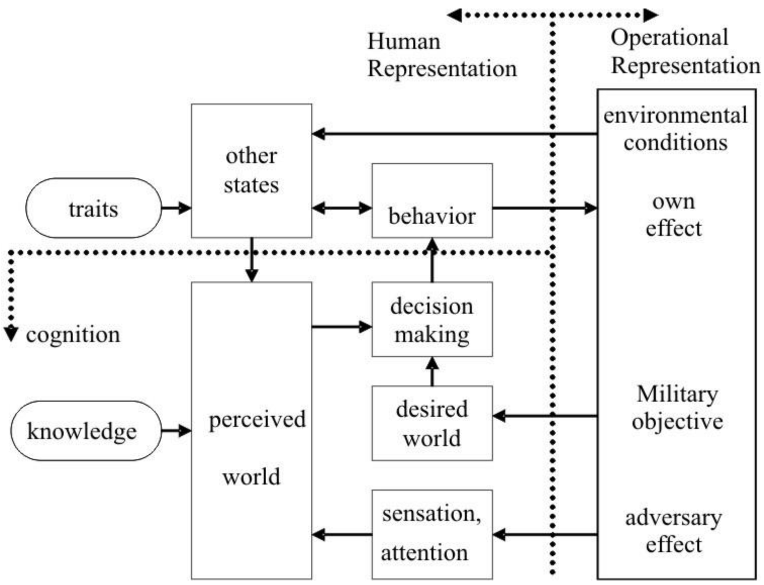

Figure 6: Scheme for implementing Human Behaviour Representation in operational studies. Cognition interacts with the physiological human representation and the physical, operational representation of the environment. Behaviour is driven by decisions, trying to make the perceived world match the desired world. Behaviour is also driven by internal urges, stemming from stressors, however. Behaviour affects the operational effect that is achieved, lifting it to the level of the problem context, in reality a battle of the will, if not a fight.

Whatever the theoretical foundation, the model must keep track of the factors that will guide the selection from among alternative behaviours and must select one for execution at the required points. In Figure 6 one concept of a dynamic task network is illustrated. The tasking comes from the mission *context*, with a presumed, default starting order of task execution. The priority ultimately associated with a task depends on *external environmental factors and internal HF* (such as traits, perceived cost and benefits, history of goal activation, etc.). Alternative goals as moderated by the *state* of the human may be considered according to the theory or approach being used for decision making.

4.0 OPERATIONAL PERFORMANCE

The question "What is operational performance?" is of key importance in group tasks. As a concept, operational performance seems to be an ordinal measure of performance that can be compared between operations, maybe done with different equipment, different methods, or different training. No such unique measure exists. In the effects based context, the only success factor is achieving the operational aim. Military exercises are evaluated by judges and they typically state their evaluation in terms such as "It took longer than usual to execute the exercise, but it was completed before dawn and there was no loss of aim." This statement clarifies that the operation succeeded because the aim was met, despite taking longer than usual, but inference is that the duration was still acceptable. The aim may have been specified in more detail regarding quantity (number of weapons confiscated), quality (suspects missed at the road block), or level (rank of officer arrested in a humanitarian crime case). An operation to collect weapons may be successful if the specified number of weapons has been turned in, but only modern rifles are counted and only those in functional condition. There are just two outcomes of an operation: succeed or fail.

Both success and failure are associated with costs, such as how long it took, personnel requirement, the losses expressed in blue, green and red casualties, damage done, resources used, readiness after the action and others. If the entity has no other tasks than the one just completed, it is difficult to say what the actual cost of these losses is. Damage will be repaired, casualties treated, and readiness restored. However, if the follow-up plans involve deployment of materiel that is still damaged, if the entity is weakened by casualties and still tired, if ammunition is not replenished, the losses may prevent successful further action. It is the mission (the aim at two levels higher echelon) that determines the actual cost, ranging from little for a forgiving mission to high for a critical mission.

Consequently, in a simulation, operational performance cannot be defined without involving the higher levels in the scenario. In order to avoid excessive complexity when assessing performance at a level, the aims of the higher levels are specified and fixed such that the scenario does not accept dynamics above a certain level. This may seem so common in training or modelling that it may not even be noticed any more, but is not obvious in real operations. Stepping up immediately to high levels (calling for air support, political influence on field decisions) is becoming more common because of the availability of information in current tactics.

For the definition of metrics see Appendix 1.

Figure 7 helps to clarify operational performance in an effect based operation (EBO). This context has become relevant when it was understood that the way the operation would be carried out could best be decided by the team or unit on site, who had the best information. The team was made responsible for the creation and evaluation of response options, as long as it would lead to the operational aim. It even received more weight when patrols in Iraq and Afghanistan had to response to military and civil complex situations, while reach back to the compound was flawed. Thus the plan for the course of action and its execution is made by the same unit. Figure 8 highlights that a plan is associated with an expectation of its effect and that the real outcome may be different. If so, good information is obtained to adjust the plan

until it delivers the intended effect. That would lead to operational success. The cognitive process thus forms an integrated part of a supervised operation. In particular, this scheme shows that a flawed plan that is carried out impeccably may result in disaster and that a perfect, but non-forgiving plan that is carried out deficiently too. Individual performance and operational performance are thus loosely related.

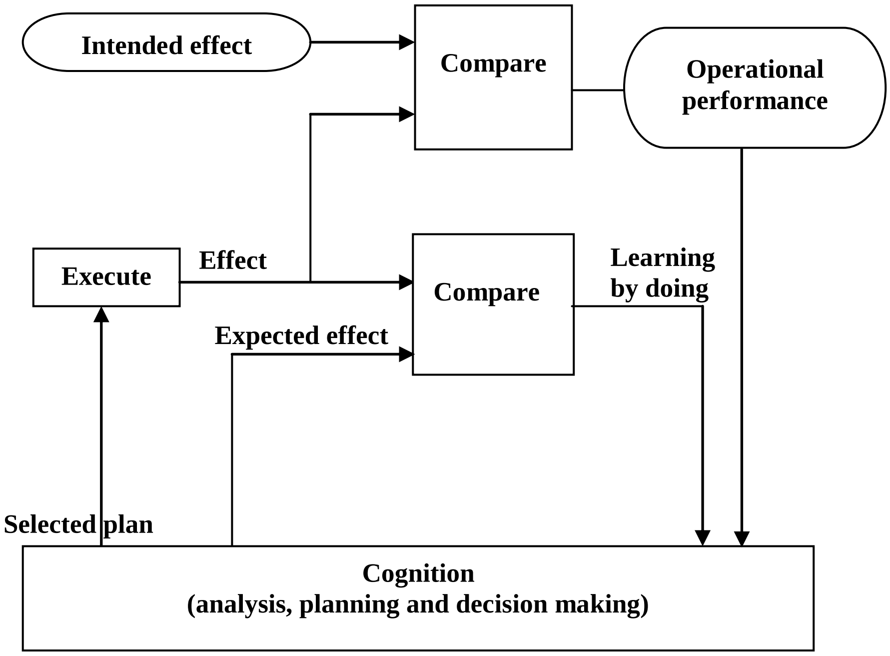

Figure 7: Goal directed behaviour.
In effects based modelling the operational performance is related to the intended effect (part of the mission). The execution may have involved costs. If the expectation is not matching the real effect, the analysis is adjusted though the feedback loop, while learning by doing. The intended effect sets the target for the operational performance.

5.0 MODELLING ARCHITECTURES

5.1 Generic Model Emerging from the HFM-128 Study

The backbone for a HF architecture in the EBO context included the three steps of Figure 1. The operation to be modelled takes place in an analysis environment, defining the units, opponents, intents, tasks, terrain, climate, means, condition of the units, etc. In many high level scenarios there is little specification of conditioning of human factors and the term analysis environment expresses that this is given appropriately. In the Human Behaviour Model both the plan and the execution is made dependent on the human states and alternative behaviours, which results in MoP's. The EBO model describes how the individuals and small teams are integrated to deliver operational performance, resulting in MoE's and cost factors. A conceptual representation of how the human model fits into the overall simulation architecture is shown in Figure 8.

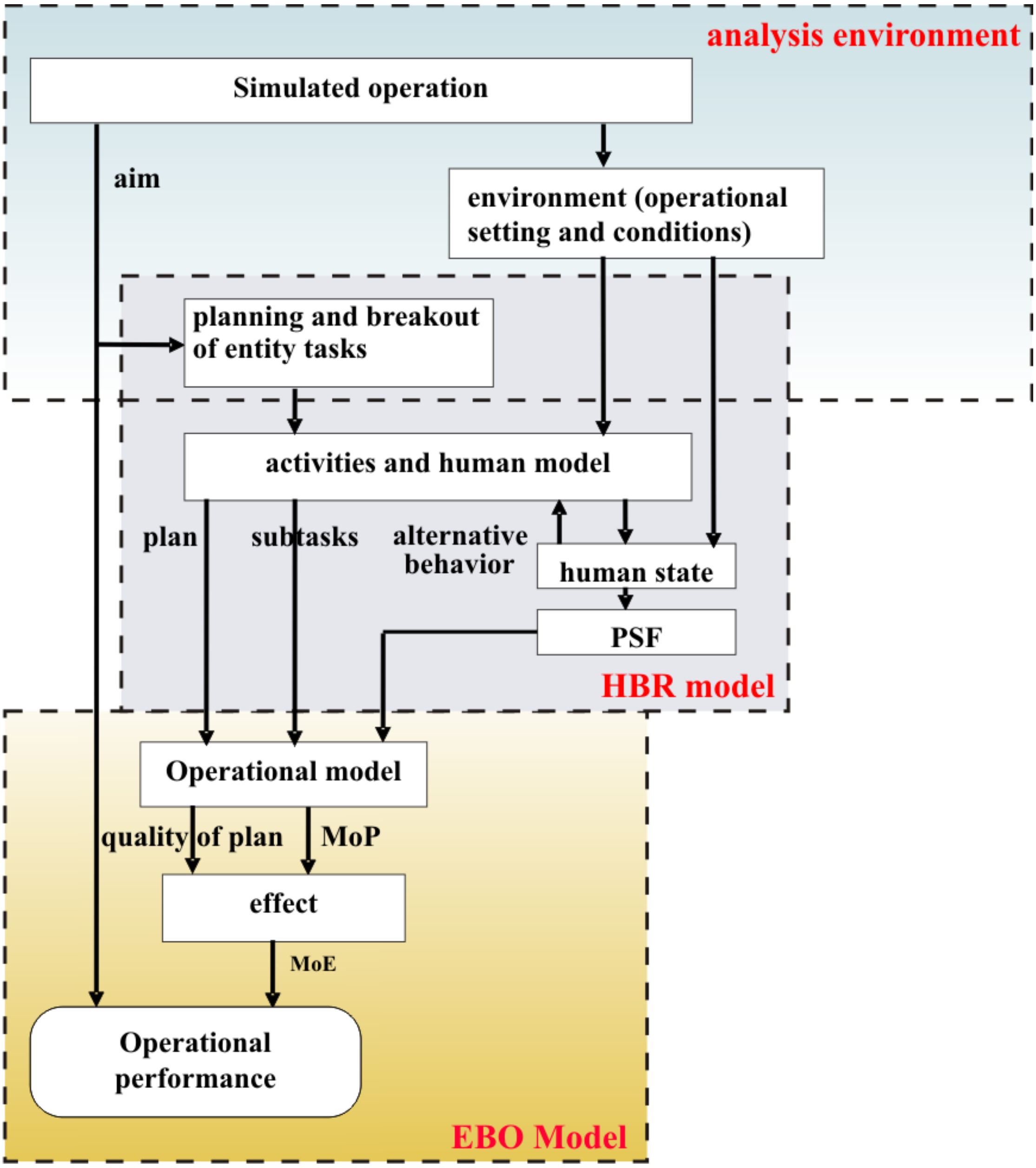

Figure 8: Synthesized model including HF in operational performance. This model is an aggregation of several previous figures to show the interactions of HBR modelling with the analysis environment for Effects Based Operations.

5.2 Available Architectures

A number of architectures are available, either as scientific tools or as a commercial product. None of these encompass the full range of modelling capacity as suggested in this paper. These architectures (see Appendix 2) have been developed for different purposes and show a variety of capabilities in the human factors, performance and operational domains. HFM 128 attempted to classify some of these architectures, to highlight the differences. The method used was to let expert users judge on the pair wise differences between sixteen chosen architectures. The specialists mentioned the following aspects as leading in their comparisons:

- Level of representation of behavior (e.g., cognitive architecture versus task network model versus scripted entity behaviors versus AI solution).
- Psychological fidelity and degree of human plausibility.
- Application area, typical use, and target market.
- Focus on cognition, perception, or behavior.
- Visualization.
- Model concepts.

The results were subjected to a Multidimensional scaling (MDS) analysis, revealing the two principal components shown in Figure 9. The architectures seem to fall apart into four clusters, one in each quadrant of the Euclidean space as shown in Figure 9.

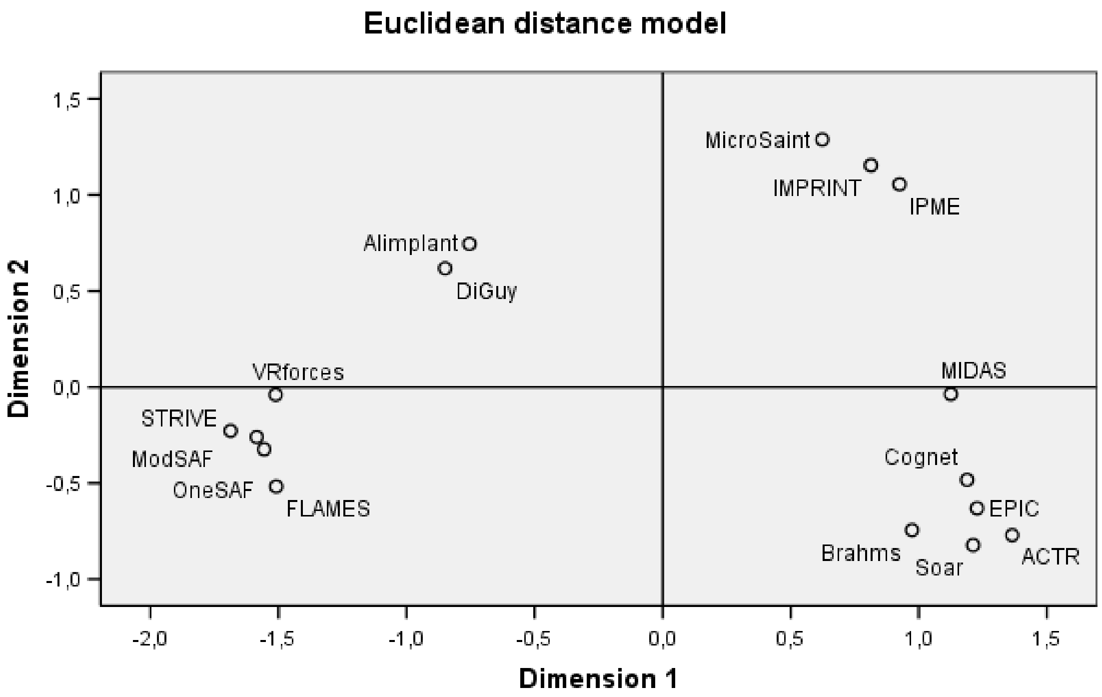

Figure 9: Factor analysis on perceived difference between modelling architectures.

The cluster in the lower left quadrant contains commercial Computer Generated Forces or scenario generation tools. The cluster in the lower right quadrant contains architectures often labeled as cognitive architectures. The upper right quadrant contains the task network architectures (although two contain micro-models of cognitive performance). The upper left quadrant contains two architectures that seem to focus on visualization of behavior or advanced digital mannequins. This indicates that even though all the 16 architectures compared here are used to model human behavior, they show rather different approaches and views upon how human behavior and cognition are represented. The different clusters also represent clearly different application areas where human behavior representation of some form is used.

Given that no modelers were expected to have extensive knowledge of all architectures, Figure 9 should not be analyzed at any finer level of detail than what has been provided above. It should be noted that for FLAMES, AI Implant, Brahms, DI Guy and STRIVE, the mean familiarity ratings were much lower than for ACT-R, IMPRINT, IPME, Soar and MicroSaint.

The two principal components in Figure 9 are not easy to interpret. Dimension 1 seems to distinguish between operational orientation (left) versus cognitive orientation (right). The meaning of dimension 2 is less clear but could be task orientation (top) versus human plausibility (bottom), at least for the two quadrants to the right.

The most significant feature in Figure 9, however, is the clustering itself. Apparently, schools of models exist that are not well integrated. Alternatively, the judges may have been familiar with their own school and perceive the distance with the other schools as large, implicitly making the same point. This could be seen as an underlining of the desire of HFM 128 to bring several stakeholders in modeling together.

6.0 RECOMMENDATIONS FROM HFM 128 TO NATO AND THE SCIENTIFIC COMMUNITY

A key recommendation is to represent human states by explicit variables that affect performance in a transparent way.

The operational military modelling community may better incorporate advances from HF and cognitive science into Human Behaviour Representations (HBRs) for constructive simulation.

Special considerations are necessary when aggregated units are used as entities in a constructive simulation. Teams and larger units have additional HF properties that do not exist at the individual level. It is recommended that the simulation be scaled by aggregating elements where necessary and incorporating the associated HF appropriate to that aggregation.

Weaknesses in HF knowledge are revealed by the demands of the latest international operations. The effects of soldier behaviour on the attitude of the local population is receiving increasing interest in the field, but the scientific knowledge in this field is still fragmented and has not reached a useful level of modelling. Advancements in this respect are encouraged.

Effects based operations have a profound impact on the way an operational problem is solved and consequently also on the requirements placed on simulation models. It is recommended that Effects Based Operations modelling exploits increased representation of cognition, in which coordinated units have their own representation of cognitive processes that capture assessment, judgment, and decision making. The choices allow for variable behaviour.

Concrete guidance regarding integrated HF modelling is provided. These are synthesized in an overall scheme and a 19-step process to guide practitioners and analysts through an HF reinforced study case, called good practice. Many steps in the procedure are not a new insight, but violated in practice.

The validation of models remains a concern. Substantial funding should be allocated for this stage of the development, particularly if data must be collected. It is important that the integrated model be tested with use cases to ensure validity, and we strongly recommend describing the validation method used when presenting the simulation results so that the audience can make an informed judgment about the suitability of the conclusions.

REFERENCES

Anderson, J. R. & Lebiere, C. (1998). The atomic components of thought.

ARL (2005). Improved Performance Research Integration Tool (IMPRINT): User's Guide Version 7 Aberdeen, MD: US Army Research Laboratory, Human Research and Engineering Directorate.

Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How simulations serve the animal-culture interface. Psychological Review, July 2010.

Bratman, M. E. (1999). Intention, Plans, and Practical Reason. CSLI Publications. ISBN 1-57586-192-5

Klein, G.A. (1998). Sources of Power: How People Make Decisions, MIT Press, Cambridge, Mass, pp. 1-30.

HFM 128 (2007). Human Behaviour Representation in Constructive Modelling. RTO-TR-HFM-128 AC/323(HFM-128)TP/250. Web publication on NATO website.

HFM 143 (2007). Human Behaviour Representation in Constructive Modelling, Toronto, RTO-MP-HFM-143 AC/323(HFM-143)TP/285, Web publication on NATO website.

HFM 209 (2010). Task group to study 'Emotions and attitudes in constructive behaviour modelling'.

Jackson, J. V. (1987). Idea for a mind. ACM SIGART Bulletin, 101, 23 - 26.

Leone, L., M. Perugini, and R.P. Bagozzi (2005). Emotions and decision making: Regulatory focus moderates the influence of anticipated emotions on action evaluations. Cognition and Emotion Volume 19, Issue 8, Pages 1175-1198.

Selfridge, O. G. (1959). Pandemonium: A Paradigm for Learning. Paper presented at the Proceedings of Symposium on the Mechanization of Thought processes.

Sheppard, B.H., J. Hartwick and P.R. Warshaw (1988). The theory of reasoned action: a meta-analysis of past research with recommendations for modifications and future research. J Cons Res, Dec 88, 325-342.

Wegner, D.M. and T. Wheatley (1999). Apparent mental causation: Sources of the experience of will. American Psychologist 54 (7), 480-492.

Zachary, W., Ryder, J. M. & Hicinbothom, J. H. (1998). Cognitive task analysis and modelling of decision making in complex environments. In J. A. Cannon-Bowers & E. Salas (Eds.), Making Decisions under Stress: Implications for Individual and Team Training (pp. 315-344). Washington, DC: American Psychological Association.

Appendices

Appendix 1. Definitions

Table A1.1: Definitions in the operational context.

Model Element	Character	Input	Output
Human Resource	Human competencies used for performing tasks	Specified in the HF literature. May be moderated by training, instrumental leveraging or environment.	Moderated human resources
Task	A familiar, basic or aggregate operational activity performed by an entity	Is generated by the plan, quantifying the task	The use of human and operational resources must be specified and the progress towards completion
State variable	A human internal parameter that is representative for the state of an HF subsystem	Human perceptual, mental or physiological processes as well as human traits	The values feed into PSFs. States may reduce available human resources.
PSF	Operation independent function describing the way a task is done given the development of particular states of the entity in the course of time	The task determines which PSFs are stimulated and how strong the stimulation is. PSFs may take states as input, but sometimes predicts performance without intervention of states.	PSFs produce quantitative task performance
Task performance	Quality of task execution related to best possible task execution	PSFs or given by the operational context as traits (training, fear, morale, etc.)	Relative achievement. Time and other cost factors are passed to the assessment of operational performance
Effect	Function describing the operational consequences of the combined tasks of the entities involved	Task performance, in the framework of the plan.	The effect created by the combined tasks and the combined cost
Operational performance	Function comparing the achieved effects to the intent	The achieved effects and the intent	Success or failure of the intent and the cost in the perspective of further actions as required by the mission

Appendix 2: Modelling architectures

References for modelling architectures

- ACT-R act-r.psy.cmu.edu
- AI-Implant www.biographictech.com
- Brahms www.agentisolutions.com
- COGNET/iGEN www.chisystems.com
- DI-Guy www.bostondynamics.com
- EPIC www.umich.edu/~bcalab/epic.html
- FLAMES www.ternion.com
- IMPRINT www.arl.army.mil/ARL-Directorates/HRED/imb/imprint/Imprint7.htm
- IPME www.maad.com/MaadWeb/products/prodma.htm
- Micro Saint www.maad.com/MaadWeb/products/prodma.htm
- MIDAS human-factors.arc.nasa.gov/dev/www-midas
- ModSAF homepage no longer available
- OneSAF www.onesaf.net
- Soar sitemaker.umich.edu/soar/home
- Strive www.cae.com

Table A2.1: Motor sub models in Integrative Architectures.

Architecture	Motor Sub models
ACT-R	Motor processors: manual, vocal, oculomotor
COGNET	Abstract, with provision for user-defined models
EPIC	Motor processors: manual, vocal, oculomotor
HOS	Eye, hand, trunk, foot, etc. micro-models
Micro Saint-based tools	Times and accuracies, plus micro-models
MIDAS (and redesigned)	Jack-animated mannequin
Neural network based tools	Sensor/motor integration, oculomotor
OMAR	Default effector models
SAMPLE	Time-delayed procedural actions
Soar	Motor processors: manual, vocal, oculomotor